THE JOURNAL OF SYMBOLIC LOGIC
Volume 21, Number 3, Sept. 1956

UNIFICATION OF UNIVERSES IN SET THEORY

W. V. QUINE

1. Standardization. The logic of quantification is conveniently developed in terms of the so-called *quantificational schemata*. The atomic quantificational schemata consist each of a schematic predicate letter ('F', 'G', etc.) attached to one or more quantifiable variables ('x', 'y', etc.). Compound quantificational schemata are built up of these atomic ones by means of truth functions and quantifiers. The business of the logic of quantification then becomes that of spotting the *valid* schemata: the ones that are true under all interpretations.

The logic of quantification is well understood and well behaved. Proof procedures in this domain are known, thanks to Gödel, to be complete. If decision procedures are not known for the whole of this part of logic, still we do know, thanks to Church, that they cannot exist. And we know how to streamline our proof procedure for quick success in the average case where proof is possible. Because of all this, there is a premium on adapting special theories, of special subject matters, to this general mold. This means representing the truths of the special theory as a class of quantificational schemata with interpreted predicate letters and a chosen universe of discourse. Once a theory has been thus regimented, it is a *standard* theory (in approximately Tarski's sense of the term).

Theories are commonly presented with multiple styles of variables: one set of letters for individuals, perhaps, another for natural numbers, another for real numbers, another for classes. In quantification theory, on the other hand, of the classical sort that I am speaking of, there is only a single style of quantifiable variable. The variables of this style — 'x', 'y', etc. — range over the whole universe of discourse.

Or, to put the matter another way: theories commonly involve a plurality of universes of discourse, but a standard theory involves only a single universe of discourse — different for different theories, often, but unique for each.

Hence a major step in standardizing a theory is the pooling of the ranges of any distinctive styles of variables, so as to come out with just one style of variable and just one universe of discourse. The way of translating the old quantifications, over plural sub-universes, into the new standard quantifications, over a single inclusive universe, is well known. Thus, suppose 'α' is a variable of distinctive style and special range, and suppose

Received March 19, 1956. This is a revised and extended version of a retiring presidential address which was read to the Association for Symbolic Logic in joint session with the Eastern Division of the American Philosophical Association at Boston, December 29, 1955.

'x' is a standard or general variable such as is to range over a single inclusive universe. Then we adopt or define a special predicate, say 'P', interpreted as true of all and only the values of the special variables of the style of 'α'. Then we paraphrase the special quantifications '$(\alpha)(\ldots\alpha\ldots)$' and '$(\exists\alpha)(\ldots\alpha\ldots)$' into general quantifications as '$(x)(Px \supset \ldots x\ldots)$' and '$(\exists x)(Px\,.\ \ldots x\ldots)$'. In short, we still limit the effective ranges of our general quantifications, as desired, by inserting appropriate restrictive clauses.

Commonly, when a theory fails of standard form through having distinctive styles of quantifiable variables with separate ranges, it will also fail of standard form in another respect: some of its predicates will make sense in application to variables of only some styles and not others. Thus, consider a theory that has one style of variables for numbers — say 'i', 'j', ... — and other styles for other purposes. The predicate '$<$', we may imagine, is admitted as meaningful between variables only of the numerical style. Thus the distinction of styles of variables enters into the syntactical criterion of meaningfulness of formulas.

Now suppose we abandon the distinction of styles, paraphrasing '$(i)(\ldots i\ldots)$' and '$(\exists i)(\ldots i\ldots)$' as '$(x)(Nx \supset \ldots x\ldots)$' and '$(\exists x)(Nx\,.\ \ldots x\ldots)$' where '$N$' is construed as true of just the numbers. Formulas of the sort '$i < j$' thereupon go over into formulas of the sort '$x < y$', but these stay within special contexts: contexts where the general variables 'x' and 'y' are bound by quantifiers whose effective range is limited by appropriately placed numberhood clauses 'Nx' and 'Ny'. We are free, if we choose, to leave matters thus. What had been a grammatical rule forbidding '$<$' between variables of other than the numerical style, has given way now to a more elaborate grammatical rule forbidding formulas of the form '$x < y$' save within quantifications of the concerned variables limited by numberhood clauses.

But, if we leave matters thus, we have not put the theory over into standard form — despite the elimination of distinctive styles of variables. In a standard theory, as in quantification theory itself, contexts meaningful for any one component formula are meaningful for any other.

So, if we are really to standardize the theory just now talked of, we must waive the elaborate grammatical rule just now contemplated, and accept formulas of the form '$x < y$' as meaningful without restriction. This means extending the originally numerical interpretation of '$<$', to include interpretation of '$x < y$' also where x or y is an individual, say, or a class, or whatever else there may be in our intended overall universe of discourse. One obvious manner of extension consists in simply construing '$x < y$' as false for all values of 'x' and 'y' other than numbers. Another course, which can sometimes be exploited for economy of primitive predicates, consists in giving '$x < y$' some supplementary interpretations useful in their own right. Thus, where x and y are spatio-temporal objects, we might construe '$x < y$' as 'x is a proper part of y'; and where x and y are classes, we might construe it as 'x is a proper subclass of y'. If '$<$' is a defined predicate rather than a primitive one, then what we commonly do is extend the interpretations of the primitive predicates in one way or another, and just copy off the old definition of '$<$' with general variables in place of numerical ones, letting the chips fall where they may.

2. Non-standard set theories. In the von Neumann set theory as formulated by Bernays [1], two sub-universes are distinguished, each with its special variables: a sub-universe of so-called sets and one of so-called classes. Two membership predicates are likewise distinguished: 'ϵ' for membership of sets in sets, and 'η' for membership of sets in classes. When we standardize this system in terms of general variables, the natural expedient is to fuse 'ϵ' and 'η' into a single membership predicate 'ϵ'. Then, where x is a set, the new '$x \epsilon y$' has the force of the old 'ϵ' or the old 'η' according as y is a set or a class. Where x is a class, the new '$x \epsilon y$' has to be given some arbitrary supplementary interpretation. One natural choice at this point would be construe '$x \epsilon y$' as false for all classes x.

But there is a different choice which proves more elegant: where x is a class, construe '$x \epsilon y$' as true if and only if y has as member a set whose members are exactly those of the class x. The gain in elegance is this: any class x and any set y whose members are the same become totally indistinguishable with respect to the primitive predicate 'ϵ':

$$(z)(z \epsilon x \,.\equiv.\, z \epsilon y) \,.\, (w)(x \epsilon w \,.\equiv.\, y \epsilon w).$$

x *becomes* y, under the strictest standard of identity formulable in terms of 'ϵ'.

We have thus moved from a non-standard theory, in which sets were not identifiable with classes, into a standard theory in which there are just classes. Certain of these classes count as sets, or, in another terminology, elements; namely, those classes which are members of classes.

In speaking of the von Neumann-Bernays theory in its original form as one in which sets were not identifiable with classes, I have not meant to suggest that the theory denied identity of sets to classes. The theory simply gave no meaning to identity as between sets and classes. Formulas containing reference to sets simply became meaningless when set variables were switched to class variables or vice versa. On the other hand the revised system automatically identifies the sets with the coextensive classes, if full agreement with respect to the primitive predicate be counted as identity; and well, of course, it may.

There is a well-known set theory that exploits sub-universes and special styles of variables much more elaborately than the von Neumann-Bernays system did; viz., Russell's theory of types, with indices on the variables to indicate the types. There is an equally well-known elimination of these special styles of variables, in Whitehead and Russell's method of typical ambiguity. This method consists in dropping indices but recognizing a formula as meaningful only if it is *stratified*, i.e., so constituted that indices could be inserted conformably with the theory of types. Now it is important to note that this way of eliminating special styles of variables is not sufficient to produce a standard form of theory. On the contrary, the benefits of simple quantification theory are by no means forthcoming. Substitution of 'x' for 'y' in a meaningful formula, e.g. '$x \in y$', can engender meaninglessness. Even two meaningful formulas, e.g. '$x \in y$' and '$y \in x$'; can engender meaninglessness by occurring as parts of one context. And there is a subtler failure: even when a formula is meaningful *and* visibly exhibits the form of a valid quantificational schema, still it is not *ipso facto* acceptable as a theorem.

Thus, consider:

$$(x)(x \in y) \supset (\exists z)(y \in z). \tag{1}$$

Superficially this is a case of the valid quantificational schema '$(x)Fxy \supset (\exists z)Fyz$', which can be proved in pure quantification theory thus:

$$\begin{aligned}(x)Fxy &\supset Fyy\\ &\supset (\exists z)Fyz.\end{aligned}$$

But the objection is that this proof cannot be carried through for the example (1) itself, because of failure of stratification at the middle step. If (1) is to be proved, it must be proved from appropriate special axioms; its superficial exemplification of a valid quantificational schema is not what clinches it. In general thus Whitehead and Russell were deprived of the use of quantification theory as an autonomous source of validities, freely applicable to set theory. They could not just say "The logic of quantification teaches . . .", and so justify a theorem of their set theory as an instance of a known theorem of quantification theory. They had rather to look up the proof in quantification theory and scrutinize its intermediate steps in each new application.

(1) would be a more forceful example if it were false. Actually, as Hilary Putnam has ingeniously proved to me, there is such an example: a stratified formula S which bears the outward form of a valid quantificational schema and yet is false under the theory of types.[1]

---

[1] S is formed by taking Hailperin's axioms [5] in conjunction with the axiom of choice, and negating the whole conjunction. Since Hailperin has proved that his axioms are adequate to NF [7], and Specker [10] has proved that the axiom of choice is inconsistent with NF, S must bear the form of a valid quantificational schema. Yet, since Hailperin's axioms come out true also under the theory of types, S is false under the theory of types (unless we are prepared to count the axiom of choice itself as false under the theory of types).

Let me here add a word in defense of NF, in relation to Specker's remarkable result.

3. Standardized theory of types. The set theory which is embodied in the theory of types *can* be got over into a system in standard form, but, as seen, the method is not that of typical ambiguity. We must use really general variables. Typically ambiguous variables are confined each to a type, albeit an unspecified type; our general variables are to range over all types. Yet the ontology to be described remains as in the theory of types: there are just individuals, classes of individuals, classes of classes of individuals, and so on, and no mixtures.

Let us take, as our point of departure for this new development, the theory of types with indices. According to the general plan outlined earlier for dispensing with special styles of variables, we shall need here a predicate 'T_0' true of all and only the individuals, also a predicate 'T_1' true of all and only the classes of individuals, and so on up; and then, for each i, we can paraphrase the special quantifications '$(x^i)(\ldots x^i \ldots)$' and '$(\exists x^i)(\ldots x^i \ldots)$' into terms of our new general variables as '$(x)(T_i x \supset \ldots x \ldots)$' and '$(\exists x)(T_i x \,.\, \ldots x \ldots)$'.

In the old theory, 'ϵ' was meaningful between variables only with consecutive ascending indices. Now it is to occur between general variables. Parallel to what was said of '$<$', we must — if we are to have a standard theory — adopt supplementary interpretations of '$x \epsilon y$' covering those objects x and y which are not of consecutively ascending types. Parallel to what was said of '$<$', the most obvious course is to construe '$x \epsilon y$' as false in all such extra cases.

But then the type of x will immediately precede that of y (symbolically: $xPTy$) if and only if

$$(\exists z)(\exists w)(x \epsilon z \,.\, z \epsilon w \,.\, y \epsilon w).$$

Hence we can define 'T_0', 'T_1', etc. in terms ultimately of 'ϵ' alone, thus:

$$T_0 x \equiv (y) - (yPTx),$$

$$T_1 x \equiv (y)(T_0 y \supset yPTx),$$

$$T_2 x \equiv (y)(T_1 y \supset yPTx),$$

and so on.

When classes (or sets; no distinction now) are treated according to the theory of types with special styles of variables, the principal axiom sche-

The failure of the axiom of choice may be felt as a counter-intuitive trait of NF, but it is no more so than other oddities known at the outset, e.g. the existence of non-Cantorian classes: classes whose members cannot all be correlated with unit classes. (Terminology from [9]; proof in [8].) By inserting the condition 'x is Cantorian' at critical points in theorems, classical results can be restored and oddities avoided; cf. [10], [9]. At the same time Specker's result enables him to prove the so-called axiom of infinity, '$\Lambda \epsilon$ Nn', long sought for NF.

mata are that of class existence:

$$(\exists y^{n+1})(x^n)(x^n \epsilon y^{n+1} .\equiv. \ldots x^n \ldots)$$

and that of extensionality:

$$(w^n)(w^n \epsilon x^{n+1} .\equiv. w^n \epsilon y^{n+1}) . x^{n+1} \epsilon z^{n+2} .\supset. y^{n+1} \epsilon z^{n+2}.$$

Translating these into terms of our general variables (with tacit understanding of universal quantifiers), we get:

(2) $$(\exists y)(T_{n+1}y . (x)(T_n x \supset: x \epsilon y . \equiv \ldots x \ldots)),$$

(3) $$T_{n+2}z . T_{n+1}x . T_{n+1}y . (w)(T_n w \supset: w \epsilon x .\equiv. w \epsilon y) . x \epsilon z .\supset. y \epsilon z.$$

But we also come to need some additional axioms about types, in order to make certain matters explicit which had been implicit in the use of indices on the variables. A typical axiom schema to this purpose is as follows:

(4) $$x \epsilon y .\supset. T_n x \equiv T_{n+1}y.$$

The predicates 'T_n', 'T_{n+1}', and 'T_{n+2}' in (2)—(4) are defined (for each specific n) in terms of just 'ϵ'. Consulting those definitions, we might try reducing (4) to a more elementary axiom or axiom schema, possibly with help of (2) and (3). I shall take a different line: keeping (4) as it is, I shall give (2) and (3) some simpler versions which are equivalent to them under (4). Viz., these:

(5) $$(\exists y)(T_{n+1}y . (x)(x \epsilon y .\equiv. T_n x . \ldots x \ldots)),$$

(6) $$T_{n+1}x . T_{n+1}y . (w)(w \epsilon x .\equiv. w \epsilon y) . x \epsilon z .\supset. y \epsilon z.$$

For, given (4), we can equate (2) and (5) as follows. By (4):

(7) $$T_n x . x \epsilon y .\equiv. T_{n+1}y . x \epsilon y.$$

But (2) is truth-functionally equivalent to:

$$(\exists y)(T_{n+1}y . (x)(T_n x . x \epsilon y .\equiv. T_n x . \ldots x \ldots)),$$

which, by (7), is equivalent in turn to:

$$(\exists y)(T_{n+1}y . (x)(T_{n+1}y . x \epsilon y .\equiv. T_n x . \ldots x \ldots))$$

and hence to (5), q.e.d.

Again, given (4), we can equate (3) and (6) as follows. '$T_{n+2}z$' is superfluous in (3), in view of (4), since '$T_{n+1}x$' and '$x \epsilon z$' are also there. Moreover the part '$T_n w \supset: w \epsilon x .\equiv. w \epsilon y$' of (3) amounts to '$T_n w . w \epsilon x .\equiv. T_n w . w \epsilon y$', and hence, by (7), to '$T_{n+1}x . w \epsilon x .\equiv. T_{n+1}y . w \epsilon y$', from which '$T_{n+1}x$' and '$T_{n+1}y$' can be dropped because already assumed in (3). So (3) is reduced to (6).

In (4)—(6) we have axiom schemata for a standardization of the theory of types. Now the question arises of dropping the clauses '$T_{n+1}x$' and '$T_{n+1}y$' from (6). Actually they are needed in (6), on two counts. One need of them is to prevent x and y from being individuals. Any individuals x and y would, for lack of members, vacuously fulfill '$(w)(w \in x \,.\!\equiv.\, w \in y)$'; hence (6) would identify all individuals with one another and with the null classes of all types, were it not for '$T_{n+1}x$' and '$T_{n+1}y$'. The other need of the clauses '$T_{n+1}x$' and '$T_{n+1}y$' in (6) is to assure sameness of type, and so prevent identifying the null classes of different types with one another; for any null classes x and y, even though of distinct types, vacuously fulfill '$(w)(w \in x \,.\!\equiv.\, w \in y)$'.

It is likewise so as to assure the existence of a null class in every class type that '$T_{n+1}y$' is needed in (5). The clause is indeed superfluous if y has members, since the type of y is then fixed by that of its members, as specified in the clause '$T_n x$'.

It is possible, however, to simplify (5) in another respect: we can change '$T_n x$' in it to '$xPTy$', and this is a simplification in that the complexity of '$xPTy$' in terms of '$\in$' does not mount with n. That the result:

(8) $$(\exists y)(T_{n+1}y \,.\, (x)(x \in y \,.\!\equiv.\, xPTy \,.\, \ldots x \ldots))$$

is equivalent to (5), given (4), is seen as follows. By (4),

$$x \in u \,.\!\supset.\, T_n x \equiv T_{n+1}u,$$
$$u \in w \,.\!\supset.\, T_{n+1}u \equiv T_{n+2}w,$$
$$y \in w \,.\!\supset.\, T_{n+1}y \equiv T_{n+2}w.$$

Therefore

$$x \in u \,.\, u \in w \,.\, y \in w \,.\!\supset.\, T_n x \equiv T_{n+1}y.$$

I.e., by definition of 'PT',

(9) $$xPTy \,.\!\supset.\, T_n x \equiv T_{n+1}y.$$

Also, since '$T_{n+1}y$' is defined as '$(x)(T_n x \supset xPTy)$',

$$T_{n+1}y \supset.\, T_n x \supset xPTy.$$

But this, with (9), implies:

$$T_{n+1}y \supset.\, T_n x \equiv xPTy,$$

from which the equivalence of (5) and (8) is evident.

Actually we can extend (8) slightly, thus:

(10) $$(\exists y)(T_n y \,.\, (x)(x \in y \,.\!\equiv.\, xPTy \,.\, \ldots x \ldots)).$$

The effect is to include the case:

(11) $$(\exists y)(T_0 y \,.\, (x)(x \in y \,.\!\equiv.\, xPTy \,.\, \ldots x \ldots)).$$

Where $T_0 y$, '$xPTy$' is by definition false for all x, and so is '$x \in y$'. Accordingly (11) is vacuous save for '$(\exists y)\ T_0 y$'; it merely affirms that there are individuals. This same existence assumption was implicit also in the classical form of the theory of types, through the acceptance of quantificational laws such as '$(x)Fx \supset (\exists x)Fx$' for individual variables. Accordingly (4), (6), and (10) constitute another acceptable standardization of the theory of types.

4. Departures. Once the theory of types has been put thus into standard form, simplifications suggest themselves which are substantive in character — i.e., which do not issue in an equivalent theory. Thus, suppose we were willing to equate the null classes of all class types. One of the two purposes of the clauses '$T_{n+1}x \,.\, T_{n+1}y$' in (6) would then lapse. The remaining purpose of those clauses, that of stipulating that x and y be classes and not individuals, can as well be served thereafter by '$-T_0x \,.\, -T_0y$'. This is a simplication, since the complexity of '$T_{n+1}x \,.\, T_{n+1}y$' (in terms of 'ϵ') mounted with n.

Equating the null classes enables us likewise to change '$T_{n+1}x$' in (5) to '$-T_0x$'. So (5)—(6) become:

$$(12) \qquad (\exists y)(-T_0y \,.\, (x)(x \epsilon y \,.\!\equiv\!.\, T_nx \,.\, \ldots x \ldots)),$$

$$(13) \qquad -T_0x \,.\, -T_0y \,.\, (w)(w \epsilon x \,.\!\equiv\!.\, w \epsilon y) \,.\, x \epsilon z \,.\!\supset\!.\, y \epsilon z.$$

(13) is no longer an axiom schema, but a single axiom.

Our equating of the null classes does not literally contradict the original theory of types, since the identities thus newly affirmed could not meaningfully be affirmed or denied in the original theory. Our move is in this respect like the equating of sets to classes which took place in our standardization of the von Neumann-Bernays theory. Nevertheless it issues in consequences contrary to an intuitive picture of the type ontology, notably:

$$(\exists x)(\exists y)(\exists z)(x \epsilon y \,.\, y \epsilon z \,.\, x \epsilon z)$$

(for, take z as $\iota\Lambda \cup \iota\iota\Lambda$). Certainly the system represented by (12)—(13) is a substantial departure from the standardized theory of types.

The clause '$-T_0y$' in (12) remains needed in order to assure the existence of Λ, which is provided by this case of (12):

$$(14) \qquad (\exists y)(-T_0y \,.\, (x)(x \epsilon y \,.\!\equiv\!.\, T_nx \,.\, -T_nx)).$$

Without '$-T_0y$', (14) would still affirm the existence of memberless objects y; but these might be merely the individuals, rather than Λ, were it not for '$-T_0y$'. The purpose of '$-T_0y$' in (12) is, in short, to keep the individuals distinct from Λ. Likewise, as seen, the purpose of '$-T_0x \,.\, -T_0y$' in (13) is to avoid identifying the individuals with Λ and one another. If we were content to limit our individuals to one, not to be distinguished from Λ (hence to none, Λ being a class), then we could drop '$-T_0x$' and '$-T_0y$' from (12) and (13). Ultimately those clauses are the price of a multiplicity of memberless objects.

But another way lies open, whereby we can drop those clauses and still keep a multiplicity of individuals. If we view each individual x not as memberless but as having itself as sole member, and indeed as identical with ιx, then we can drop '$-T_0x$' and '$-T_0y$' from (12) and (13), and still no limit on the multiplicity of individuals will supervene.

The idea is not to abolish the general distinction between ιx and x. The idea is to abolish it for individuals.

At the outset of standardizing the theory of types, we were faced with the need of supplementary interpretations of '$x \epsilon y$' covering objects x and y not of consecutive types. We chose blanket falsity. The modified course now envisaged consists in making rather this exception: where x is an individual, count '$x \epsilon x$' as true.

In this second departure from the standardized theory of types, then, (12) and (13) reduce to:

$$(15) \qquad (\exists y)(x)(x \epsilon y \,.\!\equiv\!.\, T_nx \,.\, \ldots x \ldots),$$

$$(16) \qquad (w)(w \epsilon x \,.\!\equiv\!.\, w \epsilon y) \,.\, x \epsilon z \,.\!\supset\!.\, y \epsilon z.$$

Moreover, the definition of 'T_nx' must be revised to suit the new idea of individual. Skipping '$xPTy$,' we can define 'T_0x' as '$(y)(y \epsilon x \,.\!\equiv\!.\, y = x)$' — where '$y = x$' in turn is '$(u)(y \epsilon u \,.\!\supset\!.\, x \epsilon u)$'. Then we can define '$T_{n+1}x$' as '$(y)(y \epsilon x \,.\!\supset T_ny)$'. It results that where x is an individual, T_nx for all n; where x is a class of individuals, T_nx for all $n > 0$; where x is a class of classes of individuals, T_nx for all $n > 1$; and so on. Types become cumulative. In particular, $T_n\Lambda$ for all $n > 0$.

Inasmuch as the identification of individuals with their unit classes occurred also in NF [7], it may be well to note that that theory and the present one are poles apart. NF has a class of everything; this theory does not. NF has, for every class, an unlimited complement containing everything else; this theory has unlimited complements for no classes whatever. This theory is much more like Zermelo's. I think NF has distinct advantages, and that the system in [6], as mended by Wang, is better still; but in the present paper I am pursuing other lines in order to dramatize the process of unifying the universes of theories containing multiple styles of variables.

5. Connections with Zermelo's theory. Let us now examine more closely the relation of these developments to Zermelo's theory [12]. In the latter the primitive predicate 'T_0' of individuality is in effect assumed, in addition to 'ϵ'. An extensionality principle is assumed which is the same as (13), and needs the clauses '$-T_0x$' and '$-T_0y$' for the old reason: memberlessness of individuals. Zermelo's principle of *Aussonderung*, comparable to (12), is:

$$(17) \qquad (\exists y)(-T_0y \,.\, (x)(x \epsilon y \,.\!\equiv\!.\, x \epsilon z \,.\, \ldots x \ldots)).$$

His axiom of the *power class* posits, for each z, the class y of all subclasses of z; it may be put thus:

(18) $$(\exists y)(x)(x \in y \mathrel{.}\equiv\mathrel{.} -T_0 x \mathrel{.} (w)(w \in x \mathrel{.}\supset\mathrel{.} w \in z)).$$

His remaining two axioms of class existence are those of the *union* of a class z of classes and the *pair class* of any z and w:

(19) $$(\exists y)(x)(x \in y \mathrel{.}\equiv (\exists w)(x \in w \mathrel{.} w \in z)),$$

(20) $$(\exists y)(x)(x \in y \mathrel{.}\equiv\mathrel{:} x = z \mathrel{.}\vee\mathrel{.} x = w).$$

Aside from redundancies, and more advanced assumptions such as the axiom of infinity and the axiom of choice, this is Zermelo's theory. Now it is interesting to observe that all these axioms except (20) are true equally for the ontology of the theory of types as first gently modified at the beginning of § 4 by equating the null classes. (20) conflicts, when z and w are of distinct types.

In (17—(20), 'T_n' turns up only in the form 'T_0'. Now the definition of 'T_0' in § 3 is not suitable here, since conformity of classes to the type structure is no longer assumed. Rather we must take 'T_0', for Zermelo's theory, as a primitive predicate. We could indeed define '$T_0 x$' as '$(y) - (y \in x)$', were it not for the one odd case Λ. We have the primitive predicate 'T_0' for no other purpose, finally, than segregation of Λ.

Zermelo's theory can therefore be improved by adopting from § 4 the idea of individuals as self-membered. '$T_0 x$' can then be defined again as '$(y)(y \in x \mathrel{.}\equiv\mathrel{.} y = x)$', and the extensionality principle can be reduced again from (13) to (16). Correspondingly (17) and (18) can be simplified now to:

(21) $$(\exists y)(x)(x \in y \mathrel{.}\equiv\mathrel{.} x \in z \mathrel{.} \ldots x \ldots),$$

(22) $$(\exists y)(x)(x \in y \mathrel{.}\equiv\mathrel{.} (w)(w \in x \mathrel{.}\supset\mathrel{.} w \in z)).$$

So the axioms of this variant of Zermelo's theory are (16) and (19)—(22), apart from higher ones such as choice and infinity.

Such is Zermelo's theory, modified by taking individuals as self-members. Now let us compare this with the theory which was got in § 4 by taking individuals as self-members. That was the theory represented by (15) and (16). The interesting point is that there is no longer any incompatibility whatever, not even in the axiom (20) of the pair class. For, we saw that types become cumulative; and, when types are cumulative, the class of z and w is admissible for *any* z and w. When types are cumulative, any two things are of *a* same type; any finitely many things are of a same type.

In the two standard systems of § 4, both the one with memberless individuals and the one with self-membered individuals, there was this drawback: the axiom schema, whether (12) or (15), contains the clause '$T_n x$', and this is a defined clause whose complexity mounts with n. Now by borrowing from Zermelo we can mend this. Thus, consider first the old situation where individuals are left memberless. (12), which is the keynote of that theory, offers up this special case when '$\ldots x \ldots$' is taken as '$T_0 x$' and n as 0:

$$(23) \qquad (\exists y)(-T_0 y \,.\, (x)(x \epsilon y \,.\!\equiv T_0 x)).$$

But this one case of (12) suffices, with Zermelo's (17) and (18), to yield all further cases of (12). For, by (23), there is the class of all individuals. Then, taking that class as the z of (18), we infer by (18) that there is the class of all classes of individuals; and so on. Thus an exhaustive class for each type is guaranteed. But then (17), with z taken as the exhaustive class of type n, gives (12).

We see therefore that we can supplant (12) by (17), (18), and (23), and thus avoid '$T_n x$'.

(19) is superfluous in a certain curious sense, viz. this: no existences are implied by (12) and (19) that are not implied by (12) alone — nor, therefore, by (17), (18), and (23).

Here, then, is the neatest systematization I know for the theory of the beginning of § 4 (the standardized theory of types tempered by fusion of null classes, but still with memberless individuals). The primitive predicate is 'ϵ'. 'T_0' is defined still as in § 3. The axioms are (13), (17), (18), and (23).

In parallel fashion a neater systematization is forthcoming for the later theory of § 4 — the one with cumulative types and self-membered individuals. Just as (12) was derivable from (17), (18), and (23), so (15) is derivable from (21), (22), and:

$$(24) \qquad (\exists y)(x)(x \epsilon y \,.\!\equiv T_0 x),$$

the definitions now being those at the end of § 4. So the axioms are now (16), (21), (22), and (24). In other words, this standardized cumulative type theory with self-membered individuals is exactly Zermelo's theory with self-membered individuals, except that (24) is added and (19) and (20) are dropped. Moreover, both (19) and (20) stay true this time; they are dropped only as dispensable. In substance, therefore, the cumulative theory of types with general variables and self-membered individuals comes out exactly the same as Zermelo's theory with self-membered individuals, together with the guarantee (24) of a class of all individuals.

At the middle of § 4 we contemplated, in passing, the outright repudiation of individuals. Fraenkel [3] has pointed out that Zermelo's theory can be so interpreted. When this line is adopted, the system of the preceding paragraph drops its (24) and falls in with Zermelo's altogether. On the other hand as long as we do want to provide for individuals in some sense, the treatment of them as self-membered seems to minimize the nuisance.

6. Functional logic. A motive for unifying universes, and thus using general variables instead of special styles, was *standardization* in approximately Tarski's sense: fitting a theory to general quantification theory, and so gaining all benefits of the latter. But unification of universes can still conduce to simplicity and elegance when there is no though of direct assimilation to quantification theory. A case in point as the theory of classes as based on Frege's *functional abstraction* and related devices.

Where '$\ldots\alpha\ldots$' is thought of as any singular term, such as 'α^2', the functional abstract is '$\grave{\alpha}(\ldots\alpha\ldots)$' in Frege's notation, or '$\lambda_\alpha(\ldots\alpha\ldots)$' in Church's, which is in most cases easier to print. The meaning is this: $\lambda_x(\ldots x\ldots)$ is the function which $\ldots x\ldots$ is of x. Thus $\lambda_x(x^2)$ is the square function.

Also there is functional application: $y`x$, in the Whitehead-Russell notation. Where $y$ is the square function, $y`3$ is 9.

Towards imbedding a theory of classes within such a theory of functions, Frege's first step was to treat sentences as singular terms. Sentences being meaningless in term positions to begin with, we can invent interpretations for such occurrences. Purposes of deriving a theory of classes within that of functions are served by arbitrarily treating sentences in term positions as terms designating an object $\top$ or an object $\bot$ according as true or false. What objects $\top$ and $\bot$ are does not matter; we can define the two signs irrespectively of that question — '$\top$' as short for an arbitrary true sentence, say '$\lambda_x x = \lambda_x x$', and '$\bot$' as short for a false one, say '$\lambda_x x = \lambda_y \lambda_x x$'.

Now to classes. In effect Frege identifies each class y with the function that yields $\top$ when applied to members of y and $\bot$ otherwise. Thereupon '$x \in y$' can be rendered '$y`x = \top$'.

The full notation of the classical theory of classes — not only '$\in$', as just now seen, but also quantification and the truth functions — now reduces to functional abstraction, application, and equality. For, we can define '$(x)(\ldots x\ldots)$' as '$\lambda_x(\ldots x\ldots) = \lambda_x \top$', and the negation sign as '$\bot =$'; and we can get conjunction by adapting a trick from Tarski [11]: '$(z)(x = (z`x = z`y))$' verifiably amounts to the conjunction of any sentences that are put for 'x' and 'y'.

To safeguard this theory against the logical paradoxes, one may adapt types and type indices to it as in Church [2]. Alternatively, however, we can keep to the single style of general variables. If we do this, then to avoid the paradoxes we must so frame our axioms as not to assume existence of quite all the functions that our functional abstracts purport to name. We can still count all the abstracts meaningful, but just reinterpret them, in the troublesome cases, as naming something other than what the notation suggests. Such a program would preserve grammatical simplicity at the price, perhaps, of some complexity in the axioms.

Still, if we are to define '$(x)(\ldots x\ldots)$' as '$\lambda_x(\ldots x\ldots) = \lambda_x \top$', then one abstract that must continue to name what it seems to name is '$\lambda_x T$'. Our ontology must therefore resemble in one respect that of *New foundations*, rather than those of the various theories of classes considered earlier in this paper: it must number among its classes the wholly exhaustive class.

What, finally, of $y`x$ where y is an individual? This question is analogous to the earlier question of '$x \epsilon y$' where y is an individual; and that question seemed best settled by identifying individuals with their unit classes. Now the way of settling the present question which promises parallel benefits is this: identify each individual y with its constant function $\lambda_x y$ — the function whose value is y for every argument.

The sort of logical function theory to which these reflecions and analogies point is one, then, with the following traits. Functional abstraction, predication, and equality are the primitive ideas. The variables are general and the grammar simple, but some of the abstracts are irregularly construed. Non-functions, or what would otherwise be non-functions, are identified with their own constant functions. There is, finally, a class of everything. How elegant a theory might be framed within these conditions is an open question.

REFERENCES

[1] PAUL BERNAYS, *A system of axiomatic set theory*, this JOURNAL, vol. 2 (1937), pp. 65–77. Continued in vols. 6–8, 13, 19.

[2] ALONZO CHURCH, *A formulation of the simple theory of types*, this JOURNAL, vol. 5 (1940), pp. 56–68.

[3] A. FRAENKEL, *Untersuchungen über die Grundlagen der Mengenlehre*, ***Mathematische Zeitschrift***, vol. 22 (1925), pp. 250–273.

[4] GOTTLOB FREGE, ***Grundgesetze der Arithmetik***, vol. 1, Jena, 1893.

[5] THEODORE HAILPERIN, *A set of axioms for logic*, this JOURNAL, vol. 9 (1944), pp. 1–19.

[6] W. V. QUINE, ***Mathematical logic***, revised edition. Cambridge, Mass., 1951.

[7] W. V. QUINE, *New foundations for mathematical logic*, ***American mathematical monthly***, vol. 44 (1937), pp. 70–80. Reprinted with additions in QUINE, ***From a logical point of view*** (Cambridge, Mass., 1953).

[8] W. V. QUINE, *On Cantor's theorem*, this JOURNAL, vol. 2 (1937), pp. 120–124.

[9] J. B. ROSSER, ***Logic for mathematicians***, New York, 1953.

[10] ERNST SPECKER, *The axiom of choice in Quine's new foundations for mathematical logic*, ***Proceedings of the National Academy of Sciences***, vol. 39 (1953), pp. 972–975.

[11] ALFRED TARSKI, *Sur le terme primitif dans la logistique*, ***Fundamenta mathematicae***, vol. 4 (1923), pp. 196–200.

[12] ERNST ZERMELO, *Untersuchungen über die Grundlagen der Mengenlehre*, ***Mathematische Annalen***, vol. 65 (1908), pp. 261–281.

HARVARD UNIVERSITY